Halftoning and quasi-Monte Carlo

Ken Hanson

CCS-2, Methods for Advanced Scientific Simulations
Los Alamos National Laboratory

This presentation available at
http://www.lanl.gov/home/kmh/

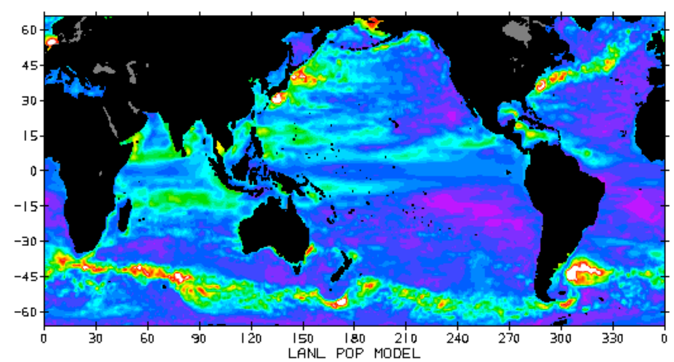
LA-UR-04-1854

Overview

- Digital halftoning – purpose and constraints
  - direct binary search (DBS) algorithm for halftoning
  - minimize cost function based on human visual system
- Quasi-Monte Carlo (QMC) – purpose, examples
- Minimum Visual Discrepancy (MVD) algorithm for points, analogous to DBS
  - examples; integration tests
- Extensions
  - higher dimensions - Voronoi, particle interaction, …
  - non-uniform sampling – adaptive, importance sampling

Validation of physics simulation codes

- Computer simulation codes
  - ► many input parameters, many output variables
  - ► very expensive to run; up to weeks on super computers
- It is important to validate codes - therefore need
  - ► to compare codes to experimental data; make inferences
  - ► use advanced methods to estimate sensitivity of simulation outputs on inputs
    - Latin square (hypercube), stratified sampling, quasi-Monte Carlo
- Examples of complex simulations
  - ► ocean and atmosphere modeling
  - ► aircraft design, etc.
  - ► casting of metals

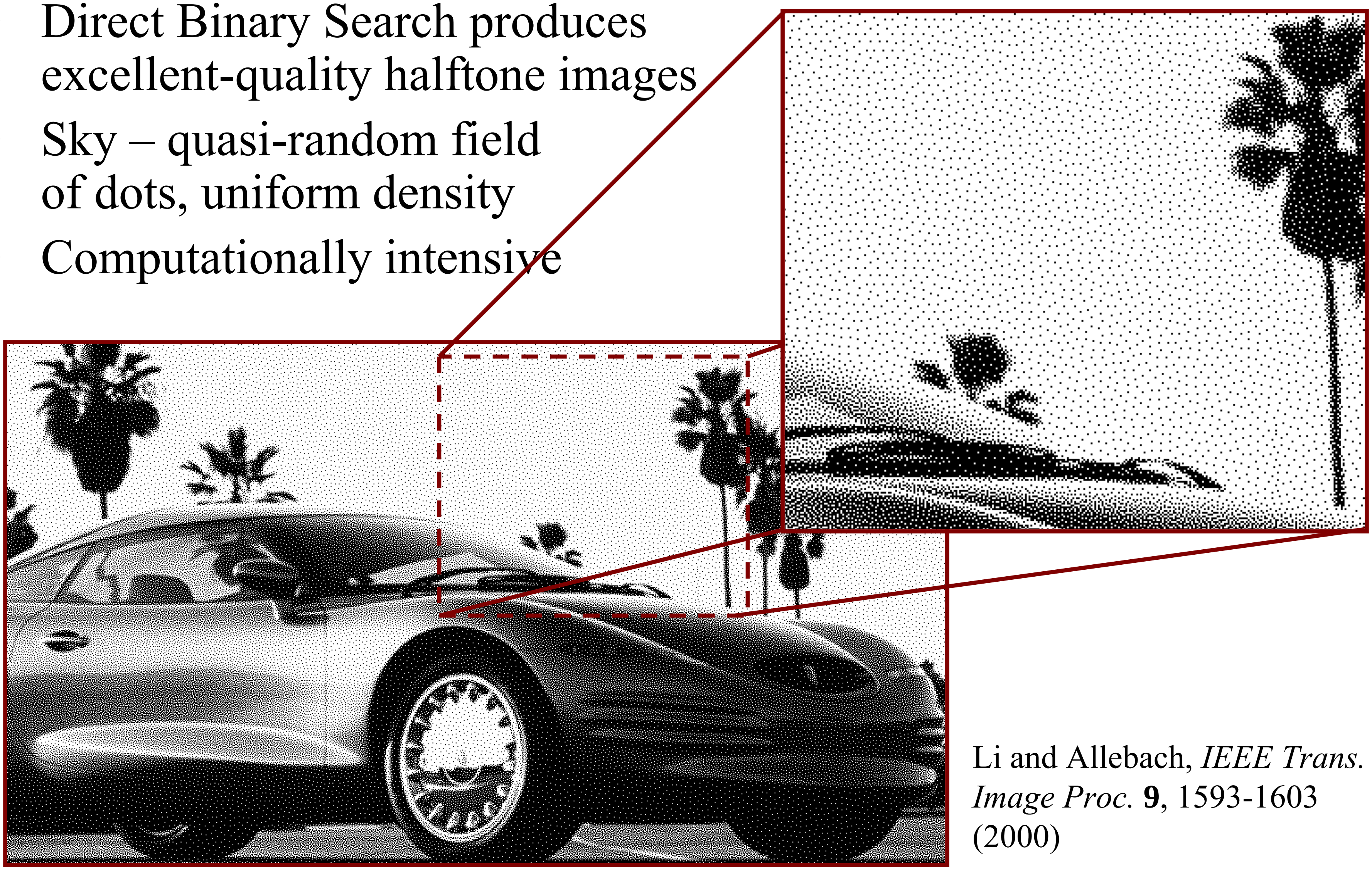


Digital halftoning techniques

- Purpose
  - render a gray-scale image by placing black dots on white background
  - make halftone rendering **look** like original gray-scale image
- Constraints
  - resolution – size and closeness of dots, number of dots
  - speed of rendering
- Various algorithmic approaches
  - error diffusion, look-up tables, blue-noise, …
  - concentrate here on Direct Binary Search

DBS example

- Direct Binary Search produces excellent-quality halftone images
- Sky – quasi-random field of dots, uniform density
- Computationally intensive



Li and Allebach, *IEEE Trans. Image Proc.* **9**, 1593-1603 (2000)

Direct Binary Search (DBS) algorithm

- Consider digital halftone image to be composed of black or white pixels
- Cost function is based on perception of two images

$$\varphi = \left|\mathbf{h} * (\mathbf{d} - \mathbf{g})\right|^2$$

  - where **d** is the dot image, **g** is the gray-scale image to be rendered, * represents convolution, and **h** is the image of the blur function of the human eye, for example, $(w^2 + r^2)^{-3/2}$
- To minimize φ
  - start with a collection of dots with average local density ~ **g**
  - iterate sequentially through all image pixels;
  - for each pixel, swap value with neighborhood pixels, or toggle its value to reduce φ

Monte Carlo integration techniques

- Purpose
  - estimate integral of a function over a specified region R in m dimensions, based on evaluations at n sample points

$$\int_R f(\boldsymbol{x})\,d\boldsymbol{x} = \frac{V_R}{n}\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} f(\boldsymbol{x}_{\boldsymbol{i}})$$

- Constraints
  - integrand not available in analytic form, but calculable
  - function evaluations may be expensive, so minimize number
- Algorithmic approaches – accuracy in terms of number of function evaluations n
  - quadrature (Simpson) – good for few dimensions; rms err ~ n^{-1}
  - Monte Carlo – useful for many dimensions; rms err ~ $n^{-1/2}$
  - quasi-Monte Carlo – reduce # of evaluations; rms err ~ n^{-1}

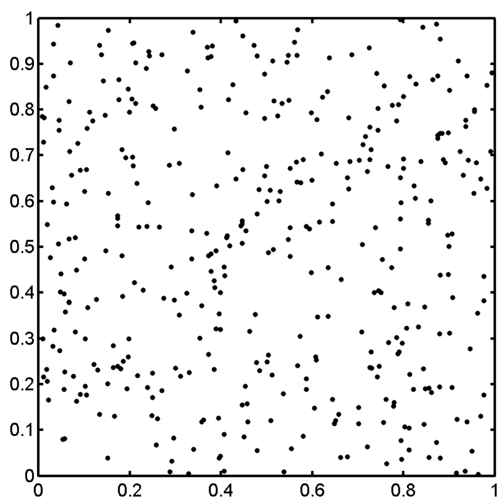
Quasi-Monte Carlo

- Purpose
  - estimate integral of a function over a specified domain in *m* dimensions
  - obtain better rate of convergence of integral estimation than seen in classic Monte Carlo
- Constraints
  - integrand function not available analytically, but calculable
  - function known (or assumed) to be reasonably well behaved
- Standard QMC approaches use low-discrepancy sequences; product space
  (Halton, Sobel, Faure, Hammersley, …)
- Propose here a new way of generating sample point sets

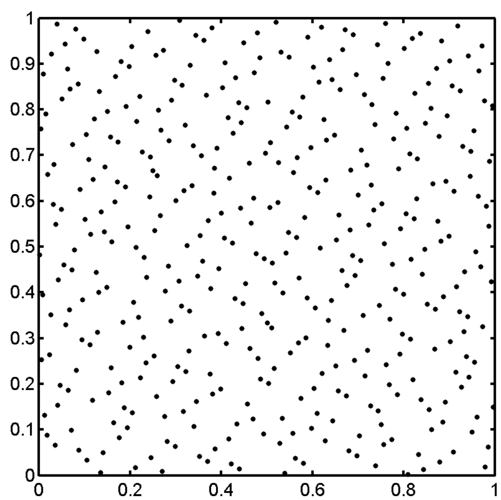
Point set examples

- Examples of different kinds of point sets
  - 400 points in each
- If quasi-MC sequences have better integration properties than random, is halftone pattern even better?

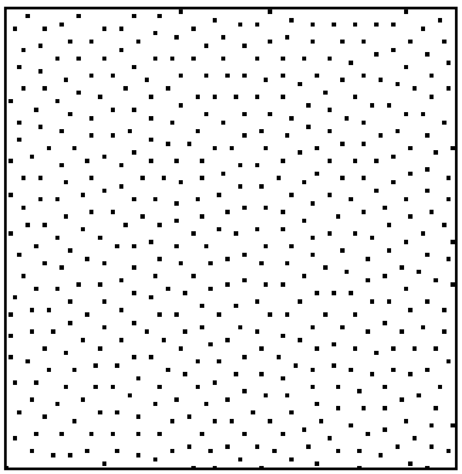
Random (independent)



Quasi-Random (Halton sequence)



Halftone (DBS sky)

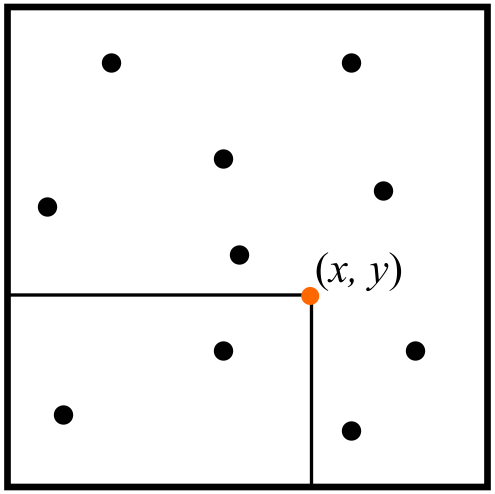


Discrepancy

- Much of QMC work is based on the discrepancy, defined for samples covering the unit square in 2D as

$$D_2 = \int_U \left[n(x, y) - A(x, y) \right]^2 \; dx dy$$

  - where integration is over unit square,
  - $n(x, y)$ is the number of points in the rectangle with opposite corners (0, 0) to (x, y), and
  - $A(x, y)$ is the area of the rectangle



- Related to upper bounds on integration error dependent on class of function
- Clearly a measure of uniformity of dot distribution

Minimum Visual Discrepancy (MVD) algorithm

Inspired by Direct Binary Search halftoning algorithm

- Start with an initial set of points
- Goal is to create uniformly distributed set of points
- Cost function is variance in blurred point image

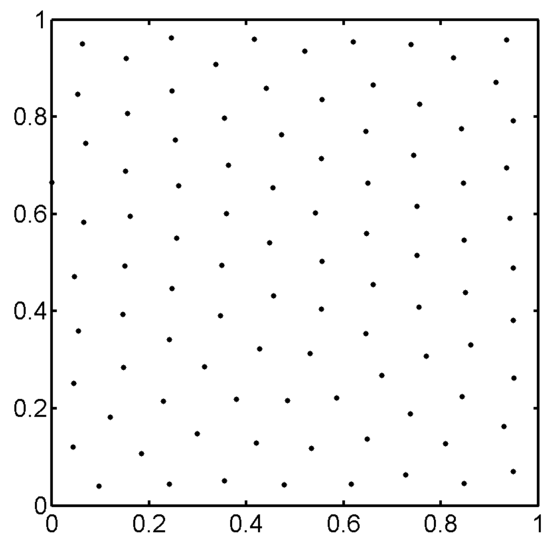
$$\psi = \mathrm{var}(\mathbf{h} * \mathbf{d})$$

  - where **d** is the point (dot) image, **h** is the blur function of the human eye, and * represents convolution
- Minimize ψ by
  - starting with some point set (random, stratified, Halton,…)
  - visit each point in random order;
  - moving each point in 8 directions, and accept move that lowers ψ the most

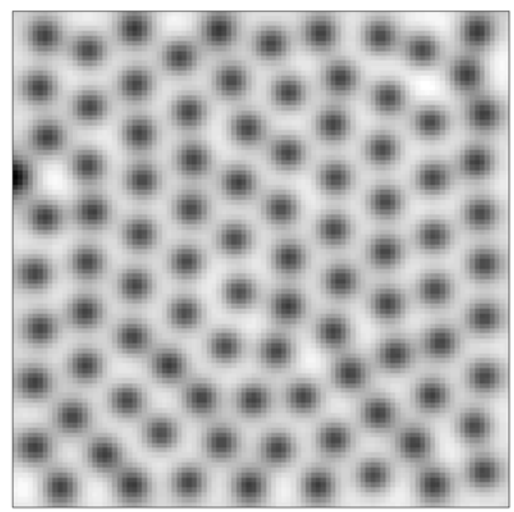
Minimum Visual Discrepancy (MVD) algorithm

- MVD result; start with 95 points from Halton sequence
- MVD objective is to minimize variance in blurred image
- Effect is to force points to be evenly distributed, or as far apart from each other as possible
- Might expect global minimum is a regular pattern

95, MVD



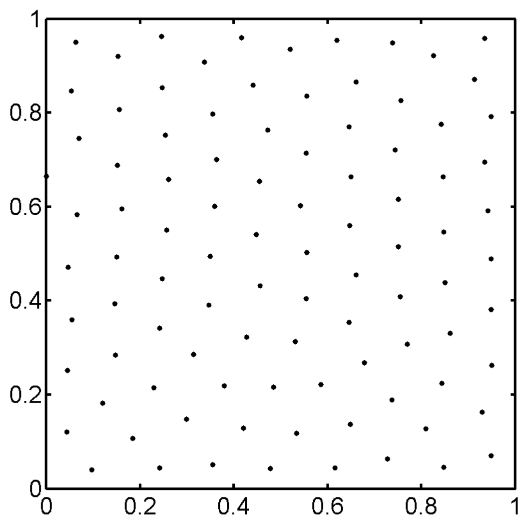
Blurred image



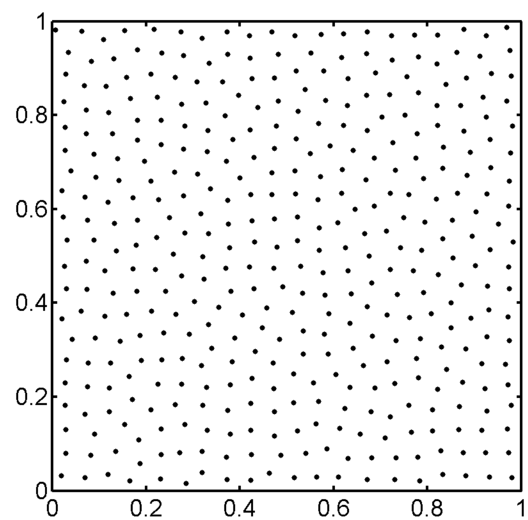
MVD results

- In each optimization, final pattern depends on initial point set
  - ► algorithm seeks local minimum, not global (as does DBS)
- Patterns somewhat resemble regular hexagonal array
  - ► similar to lattice structure in crystals or glass
  - ► however, lack long-range (coarse scale) order
  - ► best to start with point set with good long-range uniformity

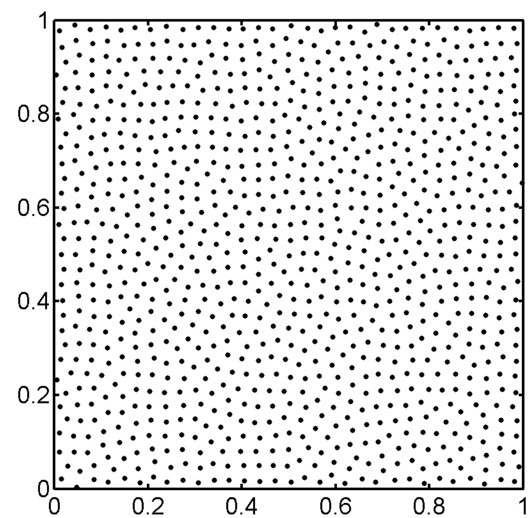
95, MVD



400, MVD



1000, MVD



Analogy to interacting particles

- Think of points as set of interacting (repulsive) particles
- Cost function is the potential

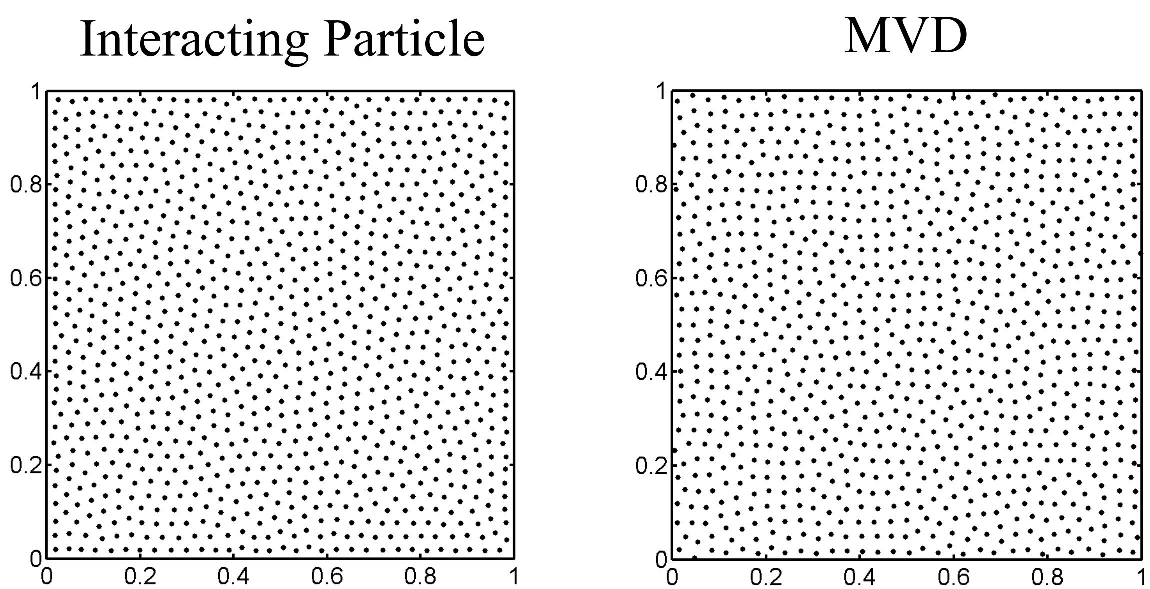
$$\psi = \sum_{\boldsymbol{i}, \boldsymbol{j} \geq \boldsymbol{i}+1} V(\boldsymbol{x_i}, \boldsymbol{x_j}) + \sum_{\boldsymbol{i}} U(\boldsymbol{x_i})$$

  - where V is a particle-particle interaction potential and U is a particle-boundary potential
  - particles are repelled by each other and from boundary
- Minimize ψ by moving particles by small steps
- This model is formally equivalent to Minimum Visual Discrepancy (V and U directly related to blur func. **h**)
- Suitable for generating point sets in high dimensions

Interacting particle approach

- Example of interacting-particle calculation
  - resulting point pattern is visually indistinguishable from MVD pattern

1000 points

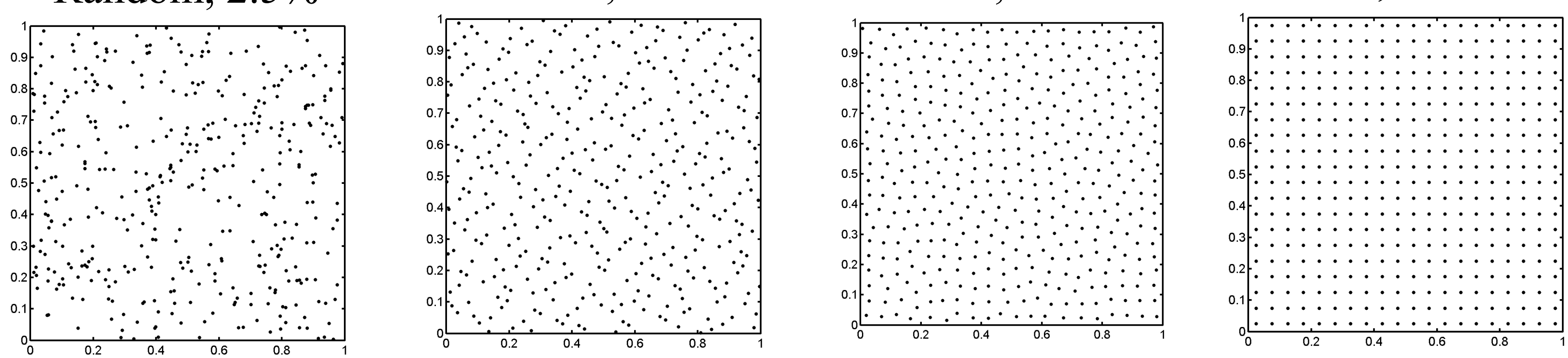


Point set examples

- Compare various kinds of point sets (400 points)
  - varying degrees of randomness and uniformity
- As the points become more uniformly distributed, the more accurate are the values of estimated integrals
- Example:

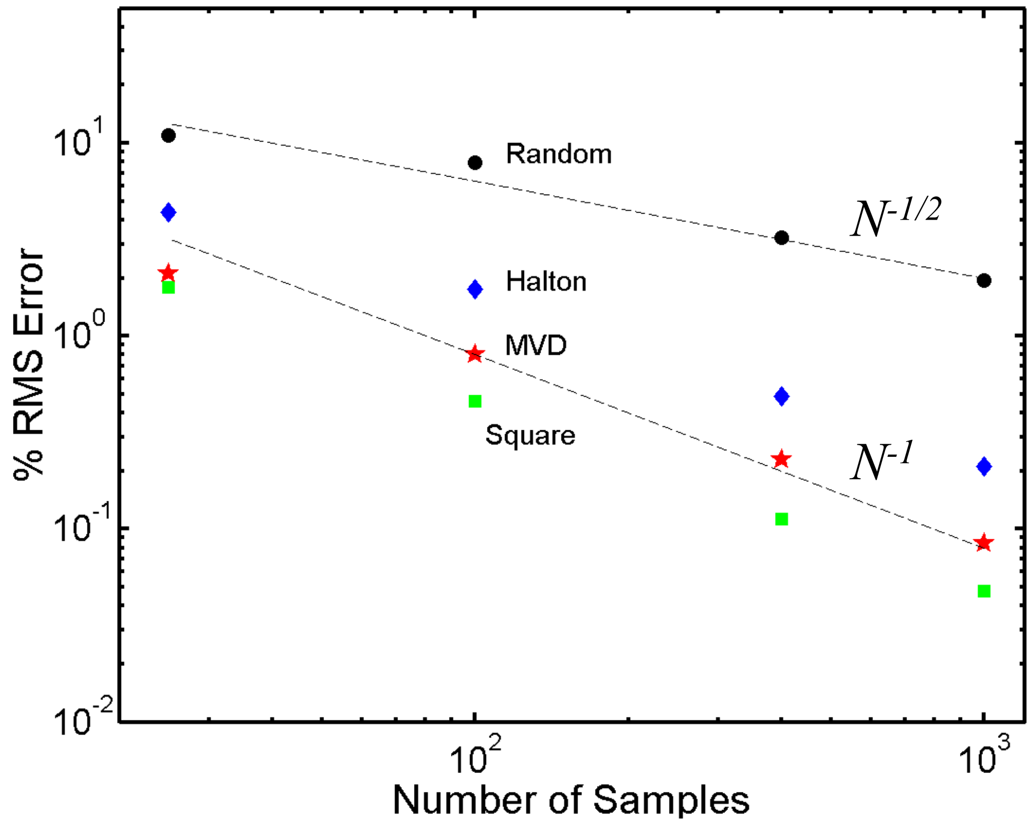
RMS relative accuracies of integral of $\text{func2} = \prod_{i} \exp\left(-2\left|\boldsymbol{x}_i - \boldsymbol{x}_i^0\right|\right); \quad 0 < \boldsymbol{x}_i^0 < 1$

Random, 2.5% Halton, 0.5% MVD, 0.14% Grid, 0.09%



More Uniform, Higher Accuracy

Integration test problems

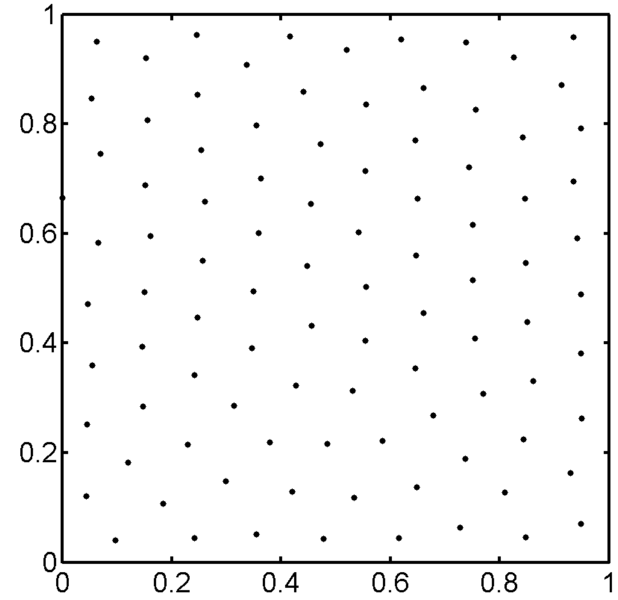


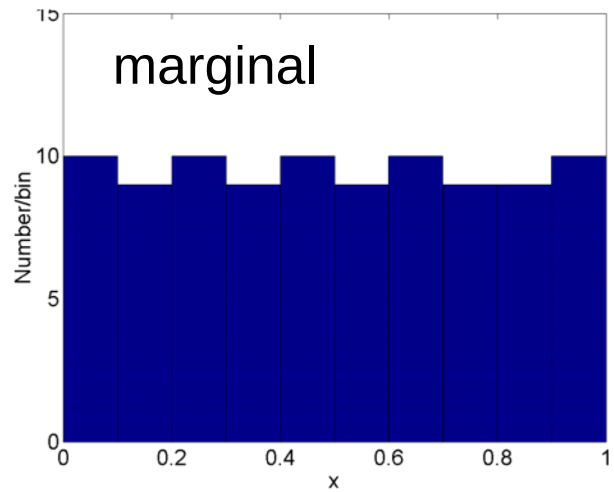
- RMS error for integral of $\text{func2} = \prod_i \exp\left(-2\left|x_i - x_i^0\right|\right); \quad 0 < x_i^0 < 1$
  - from worst to best: random, Halton, MVD, square grid
  - lines show $N^{-1/2}$ (expected for MC) and N^{-1} (expected for QMC)

Marginals for MVD points

- Sometimes desirable for projections of high dimensional point sets to sample each parameter uniformly
- Latin hypercube sampling designed to achieve this property (for specified number of points)
- Plot shows histogram of 95 MVD samples along x-axis, i.e., marginalized over y direction
- MVD points have relatively uniform marginal distributions

95 MVD points

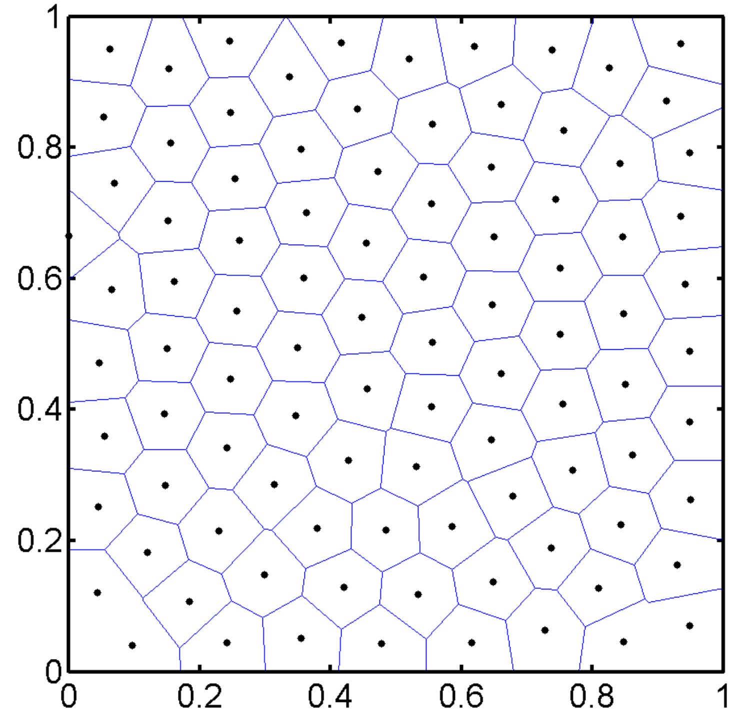




Voronoi analysis

- Voronoi diagram
  - partitions domain into polygons
  - points in *i*th polygon are closest to *i*th generating point, Z_i
- MC technique facilitates Voronoi analysis
  - randomly throw large number of points X_i into region
  - compute distance of each X_i to all generating points $\{Z_i\}$
  - sort according to closest Z_i
  - can compute A_i, radial moments, identify neighbors, …
- Extensible to high dimensions

100, MVD



Voronoi analysis can improve classic MC

- Standard MC formula

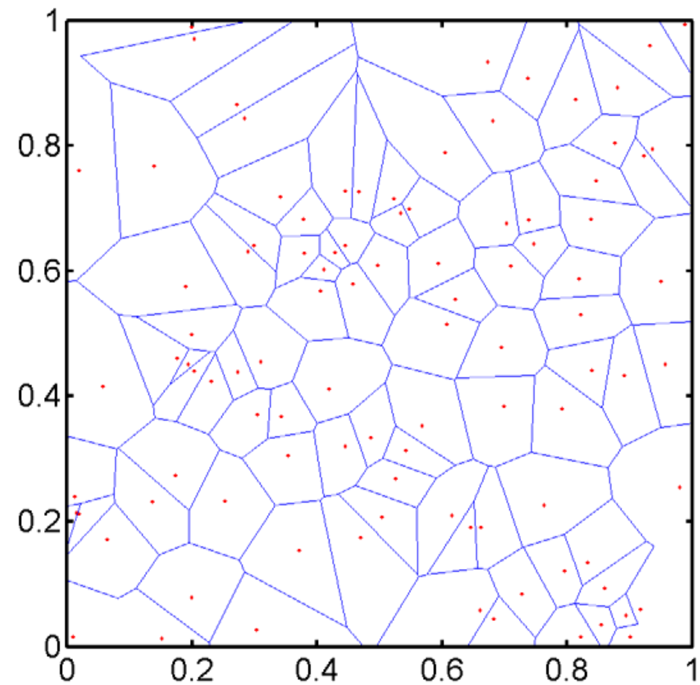
$$\int_R f(\boldsymbol{x})\,d\boldsymbol{x} = \frac{V_R}{n}\sum_{i=1}^{n} f(\boldsymbol{x}_i)$$

- Instead, use weighted average

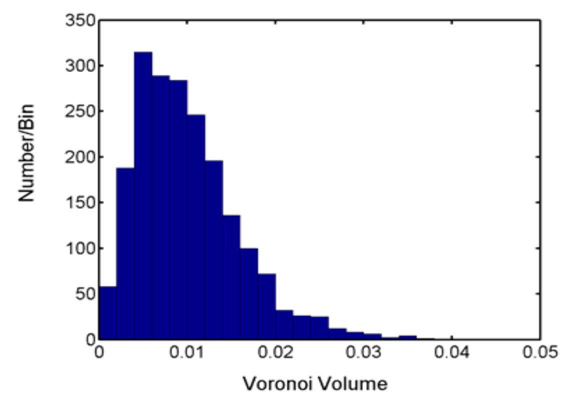
$$\int_R f(\boldsymbol{x})\,d\boldsymbol{x} = \sum_{i=1}^{n} f(\boldsymbol{x}_i)\,V_i$$

  - where V_i is the volume of Voronoi region for ith point; Riemann integr.
- Accuracy of integral estimate dramatically improved in 2D:
  - factor of 6.3 for $N = 100$ (func2)
  - factor of > 20 for $N = 1000$ (func2)
- Suitable for adaptive sampling
- Less useful in high dimensions (?)

100 random samples



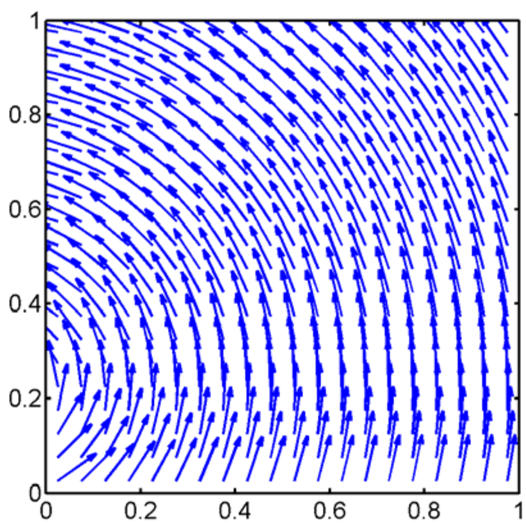
qmc91



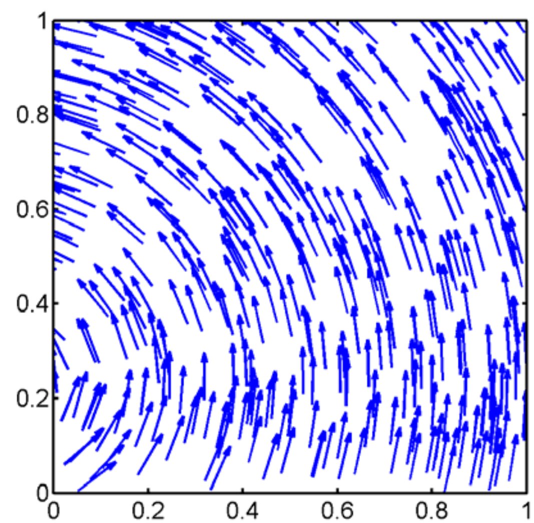
Another use - visualization of flow field

- Fluid flow often visualized as field of vectors
- Location of vector bases may be chosen as
  - square grid (typical) - regular pattern produces visual artifacts
  - random points - fewer artifacts, but nonuniform placement
  - quasi-random - fewest artifacts and uniform placement

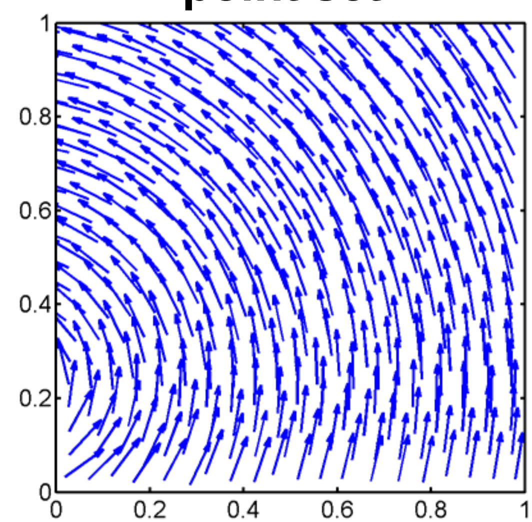
Square grid



Random points



Quasi-random (MVD) point set



Summary

- Minimum Visual Discrepancy algorithm
  - produces point sets resembling uniform halftone images
  - yields better integral estimates than standard QMC sequences
  - equivalently, can use particle interaction model
  - use MVD point sets to improve visualization of flow fields
- Extensions (using particle interaction model)
  - sampling from a specified non-uniform pdf
  - generation of optimal point sets in high dimensions
  - sequential generation of point set
    - add one point at a time, placing it at an optimal location while keeping previous points fixed
    - well suited for adaptive sampling

Comments

- Voronoi analysis
  - ► useful for determining characteristics of neighborhoods
  - ► Voronoi weighting improves accuracy of classic MC (in 2D)
    - well suited for adaptive sampling
  - ► centroidal Voronoi tessellation (Gunzberger, et al.)
- Connections to other approaches to sampling
  - ► variogram – characterizes spatial continuity (equiv. to MVD)
  - ► interpolating sampled function – kriging, local regression, etc.
  - ► Latin hyper-<u>rectangle</u> sampling (Mease et al.)
  - ► adaptive sampling (guided sequential point generation)

- Can these ideas be used MCMC for improved efficiency?

Bibliography

- K. M. Hanson, "Quasi-Monte Carlo: halftoning in high dimensions?," to appear in *Proc. SPIE* **5016** (2003)
- P. Li and J. P. Allebach, "Look-up-table based halftoning algorithm," *IEEE Trans. Image Proc.* **9**, pp. 1593-1603 (2000)
- H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, (SIAM, 1992)
- Q. Du, V. Faber, and M. Gunzburger, "Centroidal Voronoi tessellations: applications and algorithms," *SIAM Review* **41**, 637-676 (1999)

This presentation and paper available at
http://www.lanl.gov/home/kmh/